DOI:10.11931/guihaia.gxzw202306050

烟草香气相关基因 CRISPR/Cas9 编辑突变体库的构建

曾婉俐[1]，梁岗[2]，高茜[1]，李元川[3]，许力[1]，蒋佳芮[1]，许永[1]，向海英[1*]

（1. 云南中烟工业有限责任公司技术中心，云南省烟草化学重点实验室，昆明 650231；2. 中国科学院西双版纳热带植物园，热带植物资源可持续利用重点实验室，昆明 650223；3. 云南农业大学，昆明 650201）

摘　要：为探究 CRISPR/Cas9 基因编辑技术构建烟草突变体库的可行性，该研究筛选了 100 个可能参与烟草香气代谢的基因，设计相应的 100 个 sgRNA 并构建了由 100 个 CRISPR/Cas9 编辑载体组成的质粒库，获得转基因材料后分析了载体的共转化率、靶向编辑效率和脱靶编辑情况。结果表明：（1)通过农杆菌介导 100 个 sgRNA 的共转化后，在 172 个阳性转化株中检测到了其中的 77 个 sgRNA，共转化效率为 77%；（2）在 77 个携带 sgRNA 的转基因后代中，69 个 sgRNA 对目标基因进行了靶向编辑，编辑效率为 89.9%；（3）脱靶位点测序检测发现，只有一个 sgRNA 在非目标靶位点产生了脱靶编辑，表明 CRISPR/Cas9 基因编辑技术在烟草中的脱靶概率非常低。综上所述，利用 CRISPR/Cas9 载体库共转化对烟草基因进行高通量靶向编辑以构建突变体库的方法是切实可行，且该方法有共转化效率高、编辑效率高和脱靶编辑概率低等特点。

关键词：烟草，CRISPR/Cas9，烟草香气，靶向编辑，脱靶编辑

中图分类号：Q943　　　　　　　　　　**文献标识码：**A

Construction of a mutant library associated with aroma genes in tobacco using CRISPR/Cas9

ZENG Wanli[1], LIANG Gang[2], GAO Qian[1], LI Yuanchuan[3], XU li[1], JIANG Jiarui[1], XU Yong[1], XIANG Haiying[1*]

(1. *Yunnan Key Laboratory of Tobacco Chemistry, R&D Center of China Tobacco Yunnan Industrial Co., Ltd*., Kunming 650231, China; 2. *CAS Key Laboratory of Tropical Plant Resources and Sustainable Use, Xishuangbanna Tropical Botanical Garden*, *Chinese Academy of Sciences*, Kunming 650223, China; 3. *Yunnan Agricultural University* , Kunming 650201 China)

Abstract: To explore the feasibility of constructing a tobacco mutant library using the CRISPR/Cas9 gene editing technology, the study designed sgRNAs of 100 aroma related genes in tobacco, constructed a plasmid library composed of 100 corresponding CRISPR/Cas9 editing vectors, and analyzed the co-transformation rate, target editing efficiency, and off-target editing rate of transgenic offspring. The

基金项目：云南省烟草化学重点实验室开放课题（2021539200340248）；中国烟草总公司重大科技项目（110202101034（JY-11））。

第一作者：曾婉俐，博士，主要从事烟草育种研究，(E-mail)165691973@qq.com。

[*]通信作者：向海英，博士，副研究员，主要从事烟草育种研究，(E-mail)casexhy@126.com。

results were as follows: (1) After co-transformation of 100 sgRNAs mediated by *Agrobacterium tumefaciens*, 77 of them were detected in 172 positive transformation strains, with a co-transformation efficiency of 77%; (2) Among 77 transgenic offspring carrying sgRNA, 69 sgRNAs edited the target genes, with an editing efficiency of 89.6%; (3) Sequencing detection revealed that only one sgRNA produced off-target editing at a non-target site, indicating a very low probability of off-target editing of CRISPR/Cas9 in tobacco. In summary, it is feasible to construct a mutant library by the co-transformation of a CRISPR/Cas9 vector library to edit a large number of candidate target genes in tobacco. This method has the characteristics of the high co-transformation efficiency, the high editing efficiency, and the low probability of off target editing.

Keywords： tobacco, CRISPR/Cas9, aroma, on-target editing, off-target editing

CRISPR/Cas9 基因编辑是一种在基因组水平上对 DNA 序列进行定点改造的遗传操作技术。CRISPR 基因编辑系统由 Cas9 核酸酶和引导 RNA（single guide RNA,sgRNA）组成。sgRNA 的一段（～20 bp）与靶基因互补的 DNA 序列可以特异性地识别靶 DNA，而 Cas9 核酸酶对靶 DNA 进行切割，造成双链靶 DNA 的断裂，诱发内源性 DNA 的非同源末端修复机制(Jinek et al., 2012)。DNA 修复过程会引起 DNA 核苷酸的缺失或增加等错配，导致基因移码突变，实现基因敲除。CRISPR/Cas9 基因编辑系统也可以对目的基因进行特定修饰（如点突变和基因插入）。例如，在 CRISPR 基因敲除构建策略的基础上，添加计划插入修饰的 DNA 片段，就可以对特定基因插入和替换遗传修饰。

CRISPR/Cas9 基因编辑系统因其低成本、构建简单和高效等优点，已经在人类、动物和植物等生物体中被广泛利用，目前已成为基因功能研究和基因定向修饰的一个有效工具(Knott & Doudna 2018; Chen et al., 2019)。在植物中的应用表明，CRISPR/Cas9 系统不仅可高效诱导产生 DNA 突变，而且诱导的突变可稳定遗传。该技术可以在基因组水平上对 DNA 序列进行定点改造以创制优异种质资源。例如，利用 CRISPR/Cas9 基因编辑技术将普通烟草中 N'基因的抗性核心区域精准替换为野生烟中 N'alata 基因的对应区域，使得栽培烟草获得了 TMV-U1 抗性（Li et al., 2023）。利用 CRISPR 技术对 NIC 基因进行编辑产生了尼古丁含量极低的烟草新种质（Qin et al., 2021）。CRISPR/Cas9 基因编辑技术还可用于水稻和番茄等作物的快速驯化（Lemmon et al., 2018; Li et al., 2018; Yu et al., 2021），为加快新品种创制奠定了基础。CRISPR/Cas9 基因编辑技术在烟草中的成功应用（Schachtsiek & Stehle, 2019; Zhang et al., 2023），预示着该技术在着未来烟草育种中蕴藏着巨大的潜力。

烟草香味是烟叶及卷烟感官质量的重要指标之一。烟草的香气前体物主要有质体色素、西柏烷和赖百当类菇醇、酚类化合物、氨基酸和还原糖、脂类、生物碱等六个大类(汪耀富等，2006)。不同烟草品种的香气前体种类及含量烟不同，因为其内在香气前体合成相关基因的表达不同(吕婧等，2014)；也就是说，烟草香气前体种类和含量很大程度上由烟草本身的遗传因素决定。随着烟草基因组测序的完成，次生代谢产物相关合成相关酶的基因注释也日趋完善。因此，从基因层面挖掘控制烟草香气前体物的关键基因，通过遗传操作改变相关基因的表达水平可以实现定向培育特定香味的

烟草品种。利用 CRISPR-Cas9 对烟叶香味基因进行高通量靶向突变，将为优异性状新种质的创建、重要功能基因的批量挖掘以及烟草分子设计育种奠定基础。

传统烟草育种工作主要靠多亲本多组合，也就是将不同优良性状的烟草种质资源进行遗传整合，构建优质烟草新品种。但是，这种传统的杂交育种方式不仅耗费大量人力及金钱成本，而且要耗费多年时间。一些优异性状因为联锁累赘而无法整合在同一品种里，也在一定程度上限制了优异品种的创制。CRISPR/Cas9 基因编辑技术可以不依赖于遗传杂交对紧密连锁的基因进行同时编辑，可以突破传统杂交育种无法将连锁优异位点进行整合的限制，因此利用 CRISPR/Cas9 基因组编辑技术有望加速烟草香味品质改良。

尽管基因编辑技术已被应用于构建水稻、玉米等突变体库，但尚未应用于烟草基因突变体库的构建。本研究以烟草栽培品种红花大金元为材料，对其 100 个可能与烟草香味品质相关的功能基因进行靶向设计，构建 CRISPR/Cas9 载体库并进行共转化，获得了烟草香味品质相关的基因编辑突变体库。基因打靶和脱靶分析表明，利用 CRISPR-Cas9 技术构建烟草突变库是切实可行的。本研究旨验证烟草高通量基因编辑技术的可行性，为烟草基因功能研究和新种质创制提供新方法和新思路。

1 材料与方法

1.1 实验材料

本文以烤烟品种红花大金元为实验材料。

1.2 靶位点序列设计

香味相关候选基因的注释参考烟草基因组数据库 NCBI *Nicotiana tabacum* Annotation Release 100（https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Nicotiana_tabacum/100/）。利用在线工具 CRISPR-GE（http://skl.scau.edu.cn/; Xie et al., 2017）设计目标基因的特异靶位点，从候选靶位点中选择位于基因编码区前端的靶点。依据靶位点设计和克隆原则设计引物，正向引物为 5'-NNNNNNNNNNNNNNNNNNNNGTTT-3'，反向引物为 5'-NNNNNNNNNNNNNNNNNNNNAATC-3'。其中正向引物的 20 个 N 是靶位点序列，反向引物的 20 个 N 是靶位点序列的反向互补序列，引物退火后的粘性末端恰好与载体 *Bsa* I 线性化后的载体形成碱基互补配对。

1.3 CRISPR/Cas9 载体构建方法

以克隆一个基因的 CRISPR/Cas9 载体为例：（1）按照上述的引物设计原则，合成相应的正向引物和反向引物；（2）将正向引物和反向引物进行退化反应，形成双链 DNA[退火反应体系：5×Annealing Buffer for DNA Oligos 4 μL，上下游引物各 4 μL（50 μmol·μL^{-1}），Nuclease-free water 补至 20 μL。反应程序：95 ℃ 5 min，每 8 s 下降 0.1℃，直至 25 ℃（约 90 min），然后 4 ℃保存）]；（3）将退火反应得到的产物连接到经 *Bsa* I 酶切的 CRISPR/Cas9 载体 pORE-Cas9/ gRNA 上[连接体系：退火产物 2 μL，酶切产物 3 μL，10×T4 DNA Ligase Buffer 2 μL，T4 DNA Ligase（400 U·μL^{-1}）

1 μL，无菌水补至 20 μL]；（4）连接产物转化 DH5α受态细胞，利用菌落 PCR 筛选阳性克隆（上游引物为 U6-F：5'-GATCTCCCAGTCACGACGTT-3'，下游引物为目的基因上靶位点的反向互补序列）；（5）扩繁阳性菌斑，提取质粒，进行测序验证。

1.4 遗传转化

将 100 个质粒的浓度均调整为 l00 ng/μL 的，每个取 1 μL，混匀，然后转化农杆菌 GV1301，涂布在含有卡那霉素和庆大霉素的 YEB 培养基上，黑暗条件下 28 ℃培养 3 d。用液体 YEB 培养基洗脱全部菌落，-80 ℃冻存，作为后期转化烟草的 CRISPR-Cas9 编辑菌种。烟草种子播种于育苗培养皿中育苗，待长到 4 片叶子（约 30 d），便可将其移入培养瓶中，于 26 ℃、16 h 光/8 h 暗条件继续培养 30 d，备用。在超净工作台中，用打孔器将叶片打成直径 5mm 大小的圆形叶盘。用 40 mL MS 液体培养基将 YEB 培养基活化的 CRISPR-Cas9 编辑菌种悬浮成菌液（OD600: 0.6~0.8）置于 50 mL 离心管内。将叶盘转移到盛有菌液的 50 mL 离心管内，浸泡侵染 10 min。之后平铺于 MS 固体培养基上，28 ℃，黑暗， 共培养 3 d。之后将叶盘下表面接触培养基，放置于含 NAA（0.02 $\mu g \cdot mL^{-1}$）、6-BA（0.5 $\mu g \cdot mL^{-1}$）、羧苄青霉素（500 $\mu g \cdot mL^{-1}$）、卡那霉素（100 $\mu g \cdot mL^{-1}$）的 MS 培养基上，在（28 ℃，光照 16 h）/（25 ℃，黑暗 8 h）条件下培养；至长出分化芽（约 2 个月后），切下分化芽（1~2 cm）插入含有羧苄青霉素和卡那霉素的 MS 培养基上进行生根培养，分化芽约 10 d 后出根。

1.5 编辑素材的鉴定

为每个烟草单株编号，每个单株取少许叶片提取 DNA。用公共引物（正向：CTTCAAAAGTCCCACATCGCTTAG；反向：TGCAGGACTAGTGGATCAGC）进行 PCR，并用正向引物对产物进行测序。通过序列比对，提取相应单株的靶位点序列。根据靶点序列匹配相应的靶基因，并在靶位点两侧设计靶基因的特异性引物，进行 PCR 扩增，对 PCR 产物测序。通过序列比对分析相应单株的靶位点是否被编辑。

2 结果与分析

2.1 靶位点设计

根据烟草基因组的功能注释以及文献调研，本研究选择了 100 个可能参与烟草香气前体物质合成相关的基因（图 1），它们编码的蛋白参与次生代谢产物合成，离子吸收转运和信号转导等生物学过程。CRISPR/Cas9 编辑技术可用来对特定基因进行编辑以创制功能缺失突变体。CRISPR/Cas9 编辑系统由 Cas9 蛋白和 sgRNA（single guide RNA）组成，其中 Cas9 负责剪切 DNA，而 sgRNA 前端 20 个碱基通过碱基互补配对的方式识别靶位点。靶位点的 3'端还应包含一个由 3 个碱基 NGG（N 代表 4 种碱基中的一种）组成的 PAM (Protospacer Adjacent Motif)基序。同一个基因通常有多个靶位点可以选择。为保证对特定基因的编辑特异性，本研究利用 CRISPR/Cas9 设计工具 CRISPR-GE 为每个基因设计了特异性的靶位点。为确保通过 CRISPR/Cas9 编辑获得的突变体为功能缺失型突变体，靶位点设计在目标基因靠前的外显子编码区区域（图 1）。针对同一个基因存在多个不同转录

本情况，靶位点选择在这些转录本的共有外显子编码区域。烟草属于异源四倍体植物，某些基因通常存在一个序列高度相似的同源基因。对一些存在序列高度相似同源序列的基因，设计工具有时无法设计靶向单一基因的 sgRNA（单靶点 sgRNA）。针对这种情况，本研究选择了可以同时靶向两个同源基因的 sgRNA（双靶点 sgRNA）（图 1）。由于靶位点的 GC 含量比例增加可以提高编辑效率，所以当存在多个可选靶位点时，本研究选择了 GC 含量最高的靶位点（图 1）。

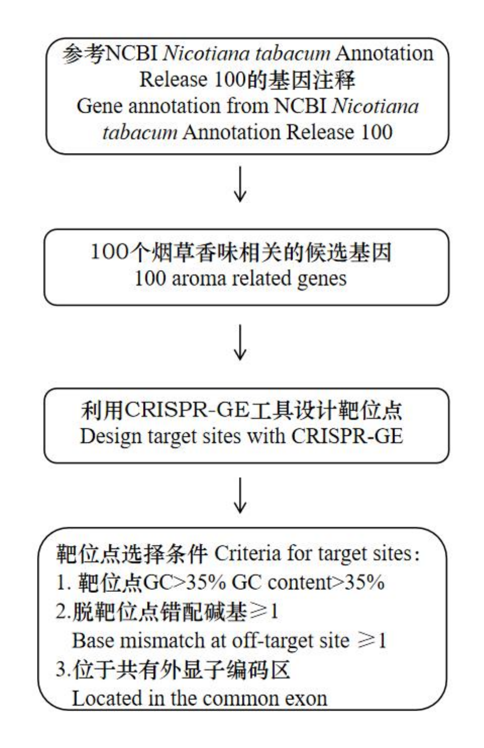


图 1 候选基因的靶位点设计

Fig. 1 Target site design of candidate genes

2.2 CRISPR/Cas9 载体库构建

在本研究所用的基因编辑载体中，烟草花叶病毒的 35S 启动子用来驱动 Cas9 基因，拟南芥的 U6-26 启动子用来驱动 sgRNA 骨架达。载体上的 sgRNA 前端预留了一段 20 bp 的插入序列，该序列包含两个 *Bsa* I 酶切位点（图 2）。载体经 *Bas* I 酶切后，产生两个特异性的粘性末端，可用来连入 20 bp 的靶位点序列。根据方法中提及的靶位点引物设计原则合成一对包含 20 bp 反向互补序列的引物，这一对引物经过退火反应后形成的双链 DNA 包含两个与载体匹配的粘性末端。由于该 DNA 的两个粘性末端恰好可以与编辑载体的粘性末端完全匹配（图 2），所以可以通过 T4DNA 连酶将片段与载体连接。连接反应后，转化大肠杆菌，抗性板筛选菌斑，PCR 扩增鉴定阳性克隆，测序正确后提取质粒。通过该流程，本研究总共构建了 100 个 CRISPR/Cas9 质粒，这些质粒按等比例混合后形成了一个 CRISPR/Cas9 载体库。

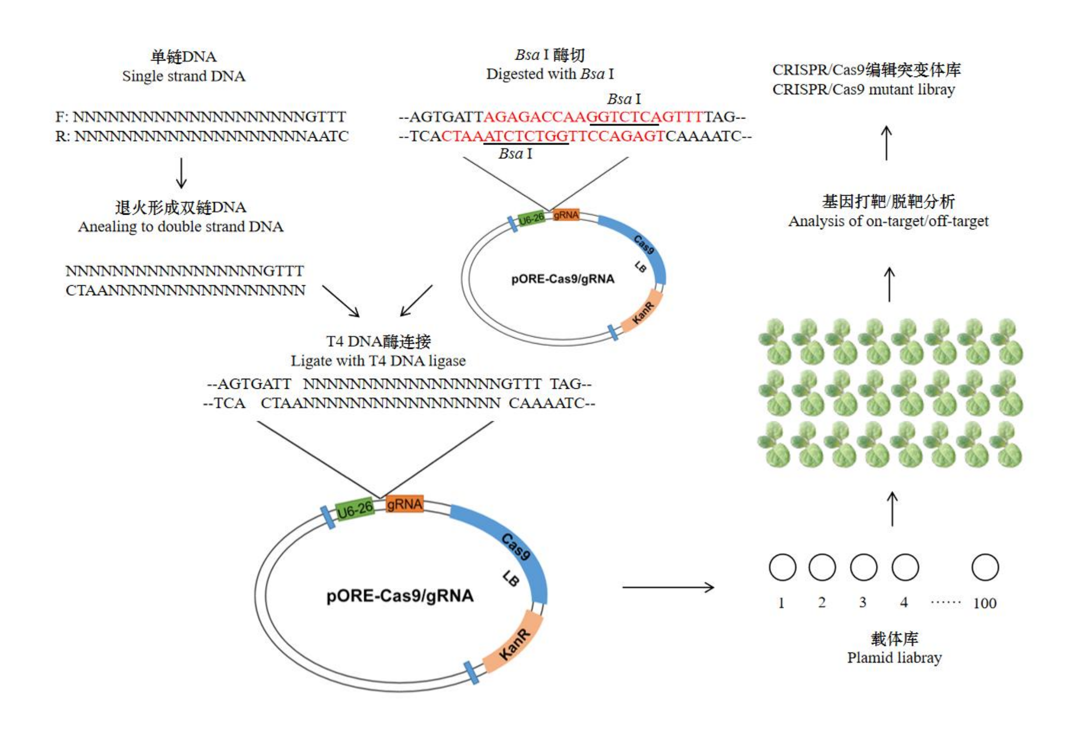


图 2 CRISPR/Cas9 编辑突变体库构建流程

Fig. 2 Schematic diagram of the CRISPR/Cas9 edited mutant library

2.3 CRISPR/Cas9 编辑分析

为了构建突变体库，本研究将上述的 CRISPR/Cas9 载体库转化农杆菌，然后通过叶盘法转化烟草，再利用抗生素筛选后获得阳性转基因烟草幼苗。由于转基因植物携带了 U6-26/sgRNA 骨架，因此可以在 U6-26 启动子区和 sgRNA 区分别设计上、下游引物并通过 PCR 测序获取阳性转基因植物的靶位点序列（图 3）。根据提取的靶位点序列找到其对应的靶位点及靶基因。在靶位点两侧设计特异性引物，PCR 扩增后测序，然后通过序列比对分析判断靶基因是否在靶位点附近被编辑。本研究从 172 个转基因植株中鉴定到 77 个不同 sgRNA 的植株。收获 T0 代阳性植株的种子后，种植 T1 代植株，然后从每个 sgRNA 株系中选择 7 个单株进行靶位点测序分析。测序结果表明，在 61 个单靶点 sgRNA 中，55 个有靶向编辑，6 个没有靶向编辑；在 16 个双靶点 sgRNA 中，13 个靶向编辑了双靶基因，1 个靶向编辑了两个靶基因的一个，2 个未靶向编辑靶（表 1）。sgRNA 的靶向编辑率为 89.6%（69/77）。

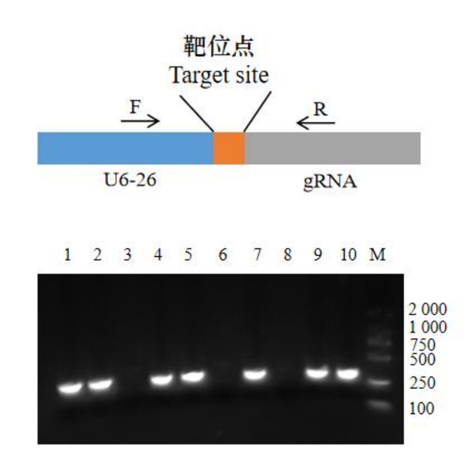


图 3 阳性植株 sgRNA 序列的扩增

Fig. 3 Acquisition of sgRNA sequences in positive transgenic plants

2.4 CRISPR/Cas9 脱靶编辑分析

为了评估编辑植株中是否存在脱靶编辑的情况，本研究利用 BLAST 软件预测了 77 个 sgRNA 的潜在脱靶位点。已有研究表明 sgRNA 与脱靶位点的碱基错配数量大于 1 个就可以阻止该位点被编辑，而 1 个碱基错配仍有较大可能会导致脱靶编辑发生（Naeem et al., 2020），因此本研究只分析了存在 1 个错配碱基的脱靶位点是否被编辑。在获得的 77 个 sgRNA 中，20 个 sgRNA 有一个脱靶位点，1 个 sgRNA 有 2 个脱靶位点（表 1）。在脱靶位点两侧设计特异性引物后进行 PCR 扩增测序，结果表明只有一个 sgRNA 的脱靶位点被编辑（图 4）。

靶位点 Target site：TTCTCCTATACCGGGCCCCTCGG

靶向基因 On-target gene：XM_016588331.1
TATTCTCCTATACCGGGCCACCTCGGTGCTG 编辑后 After editing
TATTCTCCTATACCGGGCC-CCTCGGTGCTG 编辑前 Before editing

脱靶基因 Off-target gene：XM_016581126.1
TATTCT---------------------------------------GCTG 编辑后 After editing
TATTCTCCTATACCGGGCCCCTTGGTGCTG 编辑前 Before editing

图 4 一个 sgRNA 的脱靶编辑

Fig. 4 The off-target editing of one sgRNA

表 1 突变体库基因编辑情况

Table 1 Gene editing in the mutant library

序号 No.	靶位点序列 Target site sequence	PAM	靶基因 Target gene	基因功能注释 Gene annotation	编辑 Edite	编辑率 Editing rate	脱靶基因 Off target gene	编辑 Edite	编辑率 Editing rate
1	ACCCGCATTATACAAGTCTG	AGG	XM_016631232.1	Serine carboxypeptidase	是 Yes	4/7	XM_016648616.1	否 No	0/7
2	TTCTCCTATACCGGGCCCCT	CGG	XM_016588331.1	Pullulanase	是 Yes	7/7	XM_016581126.1	是 Yes	7/7
3	ATTTTGGGGAGCCAAACCTA	GGG	XM_016658716.1	SPEAR3	是 Yes	6/7			
4	GAAGCCATTCAAGGCTGTCA	AGG	NM_001325732.2	Dihydroflavonol-4-reductase	否 No	0/7			
			NM_001325630.2	Dihydroflavonol-4-reductase	否 No	0/7			
5	GGAAGCCGGTGGTTCAGCTA	GGG	XM_016658955.1	Phenylalanine ammonia-lyase G4	是 Yes	5/7			
6	GTACATTGATTCCGTCAACA	AGG	XM_016672759.1	Senescence-specific cysteine protease	是 Yes	7/7			
7	AGACGCCGGCCACCGCCTAT	GGG	NM_001405192.1	Osmotin	是 Yes	7/7			
8	TCCTCAGCTCAACTAACCAC	TGG	XM_016615814.1	Peroxidase P7	是 Yes	7/7			
9	CTCATACATTCGGATAACAT	TGG	XM_016580175.1	CBL-interacting serine/threonine-protein kinase	否 No	0/7			
			XM_016594430.1	CBL-interacting serine/threonine-protein kinase	是 Yes	7/7			
10	GACCGTTGGGTTCTCAGTTC	CGG	XM_016600478.1	L-ascorbate peroxidase 6	是 Yes	6/7			
11	TTGTGGGTCAACAACAACTC	AGG	XM_016603279.1	Cytochrome P450	是 Yes	7/7			
12	TGTGTAGTGAACGCAGTCGC	AGG	NM_001325791.1	Tocopherol cyclase	否 No	0/7	XM_016621758.1	否 No	0/7
13	CATGTACTGATCGGGGTCGA	AGG	XM_019379535.1	Vacuolar protein sorting-associated protein	是 Yes	7/7	XM_030633095.1	否 No	0/7

14	GCCAAATTGGATGGAGCCTA	AGG	XM_016596283.1	Sugar transporter ERD6	是 Yes	7/7			
15	CACAAGAGTACCTAGCATGA	GGG	XM_016657482.1	Sucrose synthase 2	是 Yes	7/7			
			XM_016608012.1	Sucrose synthase 2	是 Yes	7/7			
16	GCATTAGGAACAGTACTGAG	AGG	XM_016655411.1	Pectate lyase	是 Yes	7/7	XM_016604443.1	否 No	0/7
17	GGCATAGAAACGGGCCGTGA	CGG	NM_001326092.1	Beta-carotene hydroxylase 2	否 No	0/7			
18	GTTCCTTCTATGGAGATATC	TGG	NM_001325221.1	Sugar transporter	是 Yes	7/7	XM_016645822.1	否 No	0/7
19	GCTGGCCTTAATCCTGATAA	TGG	NM_001325572.1	Ras-related protein RABF	是 Yes	7/7			
			XM_016650256.1	Ras-related protein RABF	是 Yes	7/7			
20	GGAGACGGCTTGTACAACGG	CGG	XM_016616459.1	Abscisic acid receptor PYL	是 Yes	7/7			
21	TGTCCAAGAGACCGTCGGGT	AGG	NM_001325338.1	Chlorophyll a-b binding protein	否 No	0/7			
			XM_016644018.1	Chlorophyll a-b binding protein	否 No	0/7			
22	CTCCTCTGCAAGACGGCCAC	GGG	XM_016635355.1	SUMO-conjugating enzyme	是 Yes	7/7			
			XM_016620991.1	SUMO-conjugating enzyme	是 Yes	7/7			
23	CTATGATAAGTCTGTACCAG	AGG	XM_016637410.1	Sucrose synthease	是 Yes	7/7	XM_016579318.1	否 No	0/7
24	GCTGCGGTTGCATGGGAAGC	CGG	XM_016647041.1	Alcohol dehydrogenase	是 Yes	7/7			
25	CTGGAGAGTTGCCTAGTGGT	TGG	XM_016602081.1	Amidase	是 Yes	6/7	XM_016629727.1	否 No	0/7
26	TTGGTTCGAGGAAGTTGCGG	AGG	XM_016645977.1	Jasmonic acid-amido synthetae	否 No	0/7	XM_016580920.1	否 No	0/7
27	AGAAGGCAGAGTAAAGACA G	AGG	XM_016642986.1	Quinolinate synthase	是 Yes	7/7			
28	GAGACCTGCAGATCTACCCG	CGG	NM_001325524.1	Reticuline oxidase	是 Yes	7/7	NM_001326242.1	否 No	0/7
29	CAAGCTGTAGGTTCCCGAGT	TGG	NM_001325814.1	Detoxification	是 Yes	6/7	NM_001325543.1	否 No	0/7

30	ATATCCATCCATTCAAGCGA	TGG	NM_016622703.1	Cytochrome P450 82C4	是 Yes	7/7			
31	ATGCGGATGAGGCTCGTCGT	AGG	XM_016658217.1	Importin subunit	是 Yes	7/7			
32	TTTGCAAGGGAGACGTCATT	TGG	NM_001325509.1	MLO-like protein	是 Yes	5/7	XM_016633095.1	否 No	0/7
33	AGCGTTCTGTGTTGCAGACA	AGG	XM_016647590.1	Reticuline oxidase	否 No	0/7			
34	TTGCCCGTAATCCAAAAGGC	TGG	NM_001325865.1	Caffeoyl-CoA O-methltransferase	是 Yes	7/7			
35	TCTGTTGTAACCGGCAAGAC	TGG	NM_001326113.1	delta12-fatty acid dehydrogenase	是 Yes	5/7			
			XM_016659636.1	delta12-fatty acid dehydrogenase	是 Yes	5/7			
36	TCGAAACAGACAAATAAACA	AGG	XM_016616941.1	ECERIFERUM	是 Yes	7/7			
37	ATTGGTGACTGGGGACGTAA	AGG	XM_016597780.1	Purple acid phosphatase	是 Yes	7/7			
38	GACCACCAGGGCACTGGCTC	TGG	NM_001324627.1	Endochitinase	是 Yes	7/7			
			XM_016653822.1	Endochitinase	是 Yes	6/7			
39	GTTACCCGACCGGTCTTACA	AGG	XM_016599368.1	Phospholipase	是 Yes	2/7			
40	CTTGGTCTGCCCGATACTGG	AGG	XM_016590747.1	Sucrose synthase	是 Yes	7/7	XM_016611295.1	否 No	0/7
41	TCCGAGGATCACCACCCAAT	GGG	XM_016640747.1	Sterol 3-beta-glucosyltran sferase	是 Yes	7/7			
42	AAATGTCCTAAGGAAAGCTG	AGG	XM_016595403.1	Sucrose synthase	是 Yes	7/7			
43	GAGGAGGGCTATTGCTGGAA	CGG	NM_001325834.1	Beta-carotene hydroxylase 2	是 Yes	7/7			
			XM_009782556.1	Beta-carotene hydroxylase 2	是 Yes	3/7			
44	GGAAGCGAATCTACGTAGGC	AGG	XR_001644690.1	Soluble starch synthase 3	是 Yes	6/7			
45	ATGGAGTTGAATTCCTCAAT	AGG	XM_016618980.1	Sucrose synthase	是 Yes	7/7			
			XM_016604255.1	Sucrose synthase	是 Yes	5/7			

46	TACAAATGTGGAAAAGCCGG	TGG	XM_016646488.1	Sterol 3-beta-glucosyltransferase	是 Yes	6/7			
47	ATACTGGGAGCAACTGAGCT	GGG	XM_016630467.1	Cysteine-rich receptor like	是 Yes	7/7			
48	GTTGTGGGGCATGAGATTAC	TGG	XM_016650499.1	Alcohol dehydrogenase	是 Yes	7/7	XM_016656110.1	否 No	0/7
49	TGGGCAAAGGAGGAGCAAAG	AGG	XM_016645712.1	Histone	是 Yes	7/7			
50	AGCTTCTTTGGCAACAGACT	TGG	XM_016624910.1	Peroxidase	是 Yes	7/7	XM_016635038.1	否 No	0/7
51	ATATCATCACGAATATATCG	CGG	XM_016612129.1	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase	是 Yes	4/7			
52	TGGGATTGAGCTGTGGGTAT	TGG	XM_016609474.1	Pectate lyase	是 Yes	7/7			
53	GCTGTCGAAAATGCATGCGG	TGG	XM_016614797.1	Peroxidase N-like	是 Yes	7/7	XM_016623966.1	否 No	0/7
54	TCGACTATAAGTGTTCCTTC	AGG	XM_016650365.1	Ferredoxin	是 Yes	7/7			
55	ATTGCCTTAGGTTACGAAGC	GGG	XM_016635292.1	Dehydrodolichyl diphosphate synthase	是 Yes	7/7			
55			XM_016628980.1	Dehydrodolichyl diphosphate synthase	是 Yes	7/7			
56	ATATGGTTCACCAAACCAAT	GGG	XM_016616473.1	Alpha carbonic anhydrase	是 Yes	7/7			
57	GCAGGGTGAGCCGACTCAAC	AGG	XM_016606799.1	V-type proton ATPase	是 Yes	7/7			
58	AACCTCTGTATGAGGGCACG	AGG	XM_016594514.1	Oxygne-evolving enhancer	是 Yes	7/7	XM_016609317.1	否 No	0/7
59	ATAAGCAGTCAACCCCACAG	CGG	XM_016622845.1	DNA replication licensing factor	是 Yes	7/7			
60	AAGTGGGATCCCAACCTCTT	TGG	XM_016634593.1	7-deoxyloganetic 8-acid glucosyltransferase	是 Yes	7/7			
61	ATTGAGCTTGCCAGAGATGT	GGG	XM_016587327.1	Diamipelargonic acid amitransferase	是 Yes	7/7	XM_016609780.1	否 No	0/7
62	CGTCGTGAGAGAAACCGTCA	AGG	XM_001325947.1	Pectin methylesterase	否 No	0/7			

63	AGCTCCCTTATCATTCCTGG	TGG	XM_016600950.1	Glutaminase	是 Yes	7/7	NM_001324687.1	否 No	0/7
							XM_016623267.1	否 No	0/7
64	CAAGCATGAAAGCTATTGTA	GGG	XM_016617986.1	Beta-gulcosidase	是 Yes	6/7			
65	TGGCGCTCTTCCCACGAAAG	AGG	XM_016619685.1	Beta-1,4-xylosyltra nsferase	否 No	0/7			
66	TATTTTGGGGAGCCAAACCT	AGG	XM_016658716.1	SPEAR3-like	是 Yes	7/7			
			XM_016638848.1	SPEAR3-like	是 Yes	5/7			
67	TTTGTCATAAACTGCGGCTG	TGG	NM_001325626.1	Nramp like	是 Yes	7/7			
			XM_016625417.1	Nramp like	是 Yes	7/7			
68	GGTTCCACTCACAGGATCTC	TGG	XM_016609851.1	DA1-related	是 Yes	7/7			
			XM_001654047.1	DA1-related	是 Yes	7/7			
69	GACTTCGTTAAACATTGCAA	TGG	XM_016581441.1	ATP carrier	是 Yes	6/7			
70	TCTACTTCATCGGCGTCAGA	AGG	XM_016578497.1	Chloride channel	是 Yes	7/7			
71	CTGCTCACTGCTCTGTACAA	AGG	XM_016642867.1	Copper transporting ATPase	是 Yes	6/7			
72	AGATTGTGCGGTGGCTAGGA	TGG	XM_016655158.1	ATPase	是 Yes	6/7			
			XM_016654239.1	ATPase	是 Yes	6/7			
73	AGGATCACCAATCTACACCT	GGG	XM_016626275.1	DA1-like	是 Yes	7/7			
			XM_016589953.1	DA1-like	是 Yes	7/7			
74	CAATGCAAGACACTGAGGAT	TGG	XM_016639678.1	ATP-dependent DNA helicase	是 Yes	7/7			
75	CAATGGTTCAAGGAGCTTCC	AGG	XM_016588741.1	NAD(P)-binding Rossmann-fold superfamily protein	是 Yes	7/7			
76	ATCAGGGAATTGACCACCAC	CGG	XM_016622691.1	Mitogen-activated kinase	是 Yes	7/7	XM_016648312.1	否 No	0/7

77	TATATCCGTGGAGACGAATA	TGG	XM_016630199.1	Receptor-like protein	是 Yes	7/7	XM_016602334.1	否 No	0/7

3 讨论

烟草的香味来源于在烟叶片中合成并累积的香味化学成分。香味相关成分的含量和种类受到相关代谢途径基因的类型及表达水平的控制。受限于烟草种质资源遗传基础方面的研究不足，烟草香味品质相关代谢物合成的基因尚未被大规模鉴定。烟草基因组测序和基因功能注释的完成为烟草香味的研究提供了新动力。本研究利用生物信息学分析获得了 100 个与烟草香味品质相关的基因，研究这些基因的功能可为后续改善烟草香味品质提供基因资源。

确定某个基因是否与烟草香味相关并对其功能进行解析很大程度上依赖于相关基因突变体的研究。基于 CRISPR/Cas9 的基因编辑技术日趋成熟，利用该技术研究和创制不同香味品质的烟草种质资源对于烟草工业的可持续发展尤为重要。利用 CRISPR-Cas9 基因编辑技术高通量构建突变体库已经应用于水稻（Lu et al., 2017; Meng et al., 2017; Chen et al., 2022）、番茄（Jacobs et al., 2017）玉米（Liu et al., 2020）和大豆（Bai et al., 2020）等二倍体植物。本研究成功将 CRISPR-Cas9 基因编辑技术应用于四倍体植物烟草的突变体库构建。最近，CRISPR-Cas9 基因编辑技术成功已应用于油菜突变体库的构建（He et al., 2023），这也是该技术首次应用于十字花科的异源四倍体植物。本研究的编辑对象烟草是茄科的异源四倍体植物，该突变体库的成功构建为在茄科多倍体植物中开展高通量靶向编辑提供了一个成功的案例。

利用农杆菌共转化获得 CRISPR-Cas9 基因编辑突变体库的方法中，共转化效率、编辑效率和脱靶编辑率是研究人员比较关注的问题。本研究的结果表明，烟草的共转化效率可以达到 77%，因此可以获得大多数载体的转基因烟草，这保证了烟草基因突变体库的高覆盖率。烟草的转基因通过愈伤组织转化实现，单次转转化的基因编辑效率高。本研究证明，共转化多基因编辑载体同样可以实现较高的基因编辑效率（89.6%）。CRISPR-Cas9 基因编辑技术潜在的脱靶作用是制约其应用的一个主要因素。本研究进行了脱靶位点分析，结果显示，77 个 sgRNA 中只有 1 个位点产生了脱靶编辑，这说明 CRISPR-Cas9 对靶位点的识别具有很强的特异性。由于本研究未进行全基因组测序分析，所以不能排除编辑植株仍然存在脱靶编辑的可能性。值得提及的是，在编辑植物中总能筛选到 Cas9 基因与编辑位点位于不同染色体上的植株，在育种实践中可以通过遗传分离和纯化选择去除转基因标签或脱靶编辑位点。

4 结论

本研究利用 CRISPR/Cas9 技术成功构建了烟草香味相关基因的突变库，表明 CRISPR/Cas9 基因编辑技术高通量编辑烟草基因可用于构建烟草突变库，该方法具有共转化效率高，靶向编辑效率高

和脱靶编辑低等特点。本研究为烟草香味功能基因的研究提供了遗传资源，为烟草香味重要性状的分子改良提供了种质资源。该技术方法并不是仅限于构建烟草香味品质相关基因的突变体库，而是可用于在烟草中构建研究人员感兴趣的任何性状的基因突变体库，这将加速烟草功能基因的研究。

参考文献：

汪耀富，高华军，刘国顺，等，2006. 不同基因型烤烟叶片致香物质含量的对比分析[J]. 中国农学通报，22（5）：117-121. [WANG YF, GAO HJ, LIU GS, et al., 2006. Study on contents of aroma constituents of different genotypes in flue-cured tobacco leaves [J]. Chin Agric Sci Bull, 22(5): 117-121.]

吕婧，刘贯山，晁江涛，等，2014. 影响烟草香气物质合成代谢途径关键酶基因的研究进展. 中国农学通报，30（4）：49-57.[LÜ J, LIU GS, CAO JT, et al., 2014. Research advances in key enzyme genes of tobacco aroma compounds synthesis metabolism [J]. Chin Agric Sci Bull, 30(4): 49-57.]

BAI M, YUAN J, KUANG H, et al. 2020. Generation of a multiplex mutagenesis population via pooled CRISPR-Cas9 in soya bean [J]. Plant Biotechnol J, 18(3): 721–731.

CHEN K, WANG Y, ZHANG R, et al., 2019. CRISPR/Cas genome editing and precision plant breeding in agriculture [J]. Ann Rev Plant Biol, 70: 667–697.

CHEN K, KE R, DU M, et al. 2022. A FLASH pipeline for arrayed CRISPR library construction and the gene function discovery of rice receptor-like kinases [J]. Mol Plant, 15(2): 243–257.

HE J, ZAHNG K, YAN S, et al., 2023. Genome-scale targeted mutagenesis in Brassica napus using a pooled CRISPR library [J]. Genom Res, 33(5):798-809.

JACOBS TB, ZHANG N, PATEL D, et al., 2017. Generation of a collection of mutant tomato lines using pooled CRISPR libraries [J]. Plant Physiol, 174(4): 2023–2037.

JINEK M, CHYLINSKI K, FONFARA I, et al., 2012. A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity [J]. Science, 337(6096): 816-821.

KNOTT GJ, DOUDNA JA, 2018. CRISPR-Cas guides the future of genetic engineering [J]. Science, 361(6405):866-869.

LEMMON ZH, REEM NT, DALRYMPLE J, et al., 2018. Rapid improvement of domestication traits in an orphan crop by genome editing [J]. Nat Plants, 4(10): 766-770.

LI T, YANG X, YU Y, et al., 2018. Domestication of wild tomato is accelerated by genome editing [J]. Nat Biotechnol, 36: 1160-1163.

LI Y, HUANG C, LIU Y, et al., 2023. CRISPR/Cas9-mediated seamless gene replacement in protoplasts expands the resistance spectrum to TMV-U1 strain in regenerated Nicotiana tabacum [J]. Plant Biotechnol J, DOI: 10.1111/pbi.14159.

LIU HJ, JIAN L, XU J, et al. 2020. High-throughput CRISPR/Cas9 mutagenesis streamlines trait gene identification in maize [J]. Plant Cell, 32(5): 1397–1413.

LU Y, YE X, GUO R, et al., 2017. Genome-wide targeted mutagenesis in rice using the CRISPR/Cas9 system [J]. Mole Plant, 10(9): 1242–1245.

MENG X, YU H, ZHANG Y, et al., 2017. Construction of a genome-wide mutant library in rice using CRISPR/Cas9 [J]. Mol Plant, 10(9): 1238–1241.

NAEEM M, MAJEED S, HOQUE MZ, et al., 2020. Latest Developed strategies to minimize the off-target effects in CRISPR-Cas-mediated genome editing [J]. Cells, 9(7):1608.

Qin Q, Humphry M, Gilles T, et al., 2021. NIC1 cloning and gene editing generates low-nicotine tobacco plants[J]. Plant Biotechnol J, 19(11):2150-2152.

SCHACHTSIEK J, STEHLE F, 2019. Nicotine-free, nontransgenic tobacco (Nicotiana tabacum I.) edited by CRISPR-Cas9 [J]. Plant Biotechnol J, 17(12):2228-2230.

XIE X, MA X, ZHU Q, et al., 2017. CRISPR-GE: A convenient software toolkit for CRISPR-based genome editing [J]. Mol Plant, 10(9):1246-1249.

YU H, LIN T, MENG X, et al. 2021. A route to de novo domestication of wild allotetraploid rice [J]. Cell, 184: 1156-1170.

ZHANG J, XING J, MI Q, et al., 2023. Highly efficient transgene-free genome editing in tobacco using an optimized CRISPR/Cas9 system, pOREU3TR [J]. Plant Sci, 326: 111523.